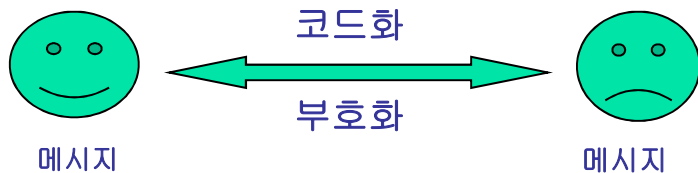
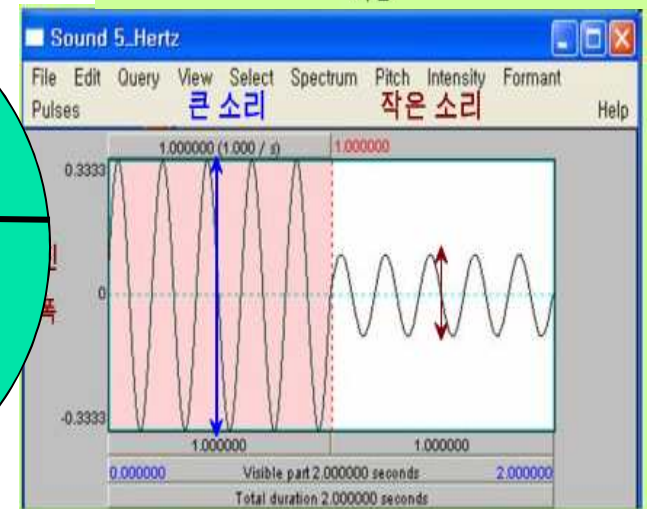
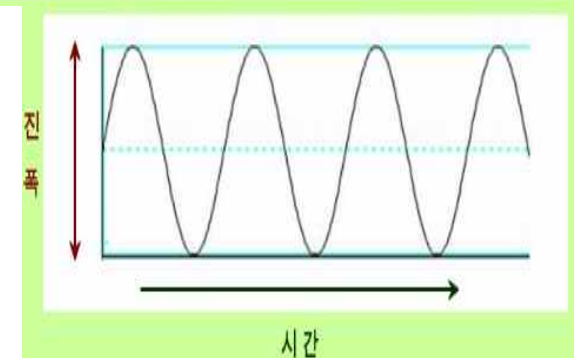
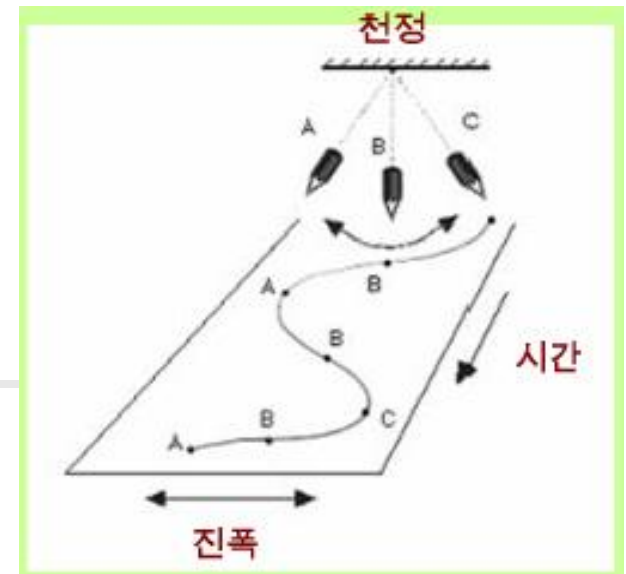
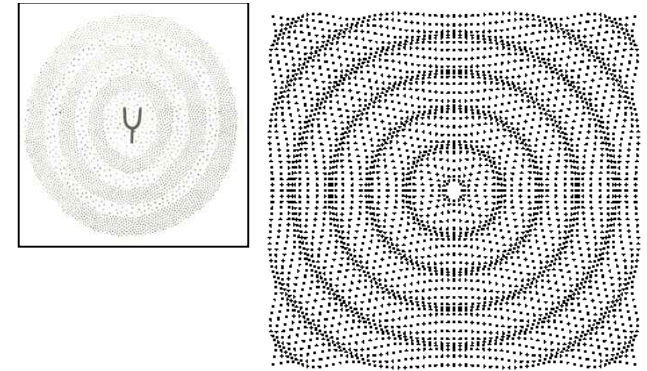
제4강. 음성산출의 이해

화 자
청 자
귀
감각신경
피드백 연결
두뇌
음파
발성근육
운동신경
두뇌
감각신경
귀
언어적 단계
생리적 단계
음향적 단계
생리적 단계
언어적 단계
www.Bandicam.co.kr
MBCHD
말하는 코끼리의 비밀

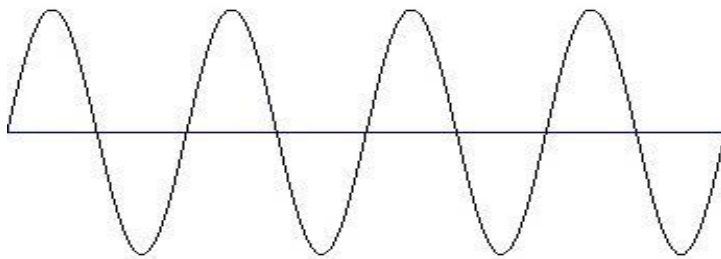
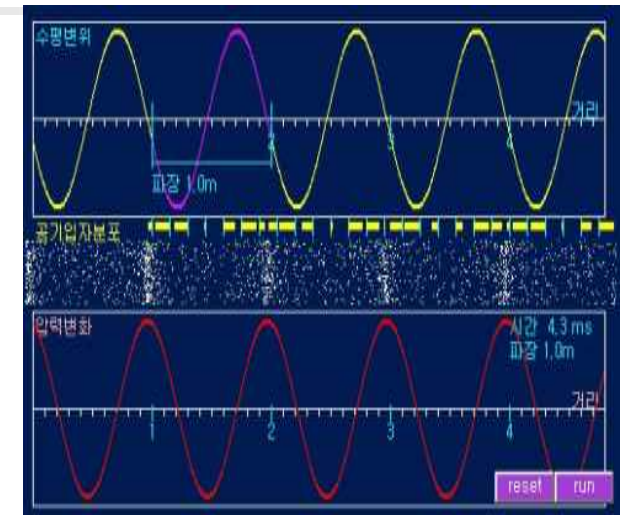
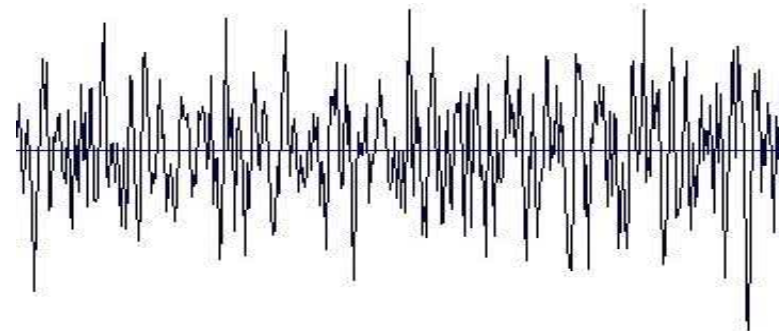
내 용

1. 말의 음향학
2. 소리의 발생과 전파
   1) 공기의 압축상태와 희박상태
   2) 주파수, 스펙트럼과 위상
   3) 공명과 포먼트
3. 말 산출과정
   1) 조음기관의 기능
   2) 모음과 자음의 산출과정
4. 말 지각과정

1. 말의 음향학

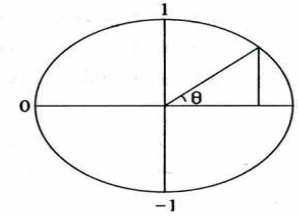
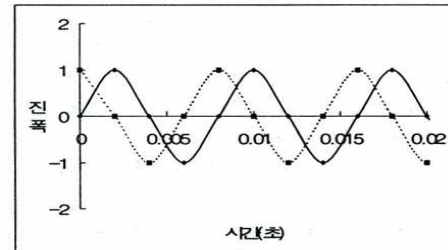
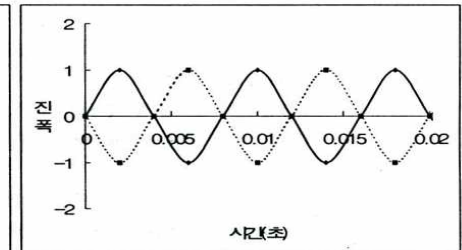
1. 말소리(**speech sound**)는 물리적 실체
   - 주파수(fundamental frequency)
   - 강도(amplitude)
   - 길이(duration)
     - => 말소리의 특성을 파악함 , 물리학적인 에너지로 정의함
2. 언어(**language**) => 의사소통체계에 관한 지식

2. 소리의 발생과 전파

1) 공기압축상태와 희박상태

1. 소리가 보인다.
2. **순음(pure tone)** : 소리 굽쇠의 떨림에 의해 만들어지는 하나의 소리
   - 공기가 압축된 상태(compression)
   - 공기가 희박한 상태(rarefaction)

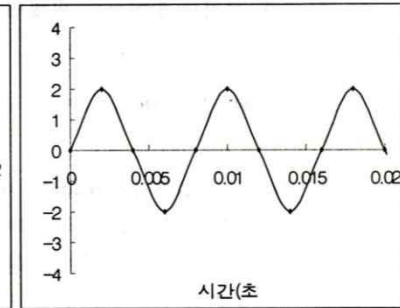
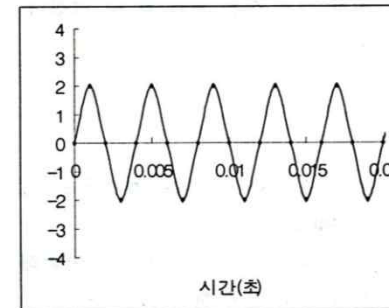
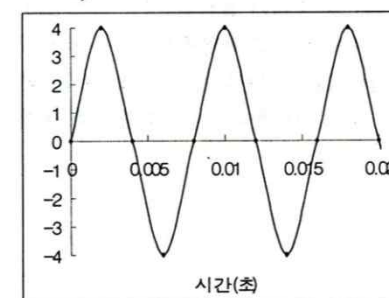
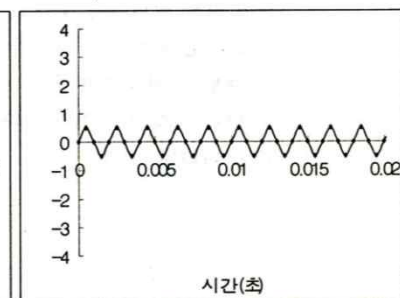
3. 소리 파형(waveform), 정현파
   - 시간에 따른 소리의 압력변화를 나타내는 그래프

(순음)

(복합음)

2) 주파수, 스펙트럼과 위상

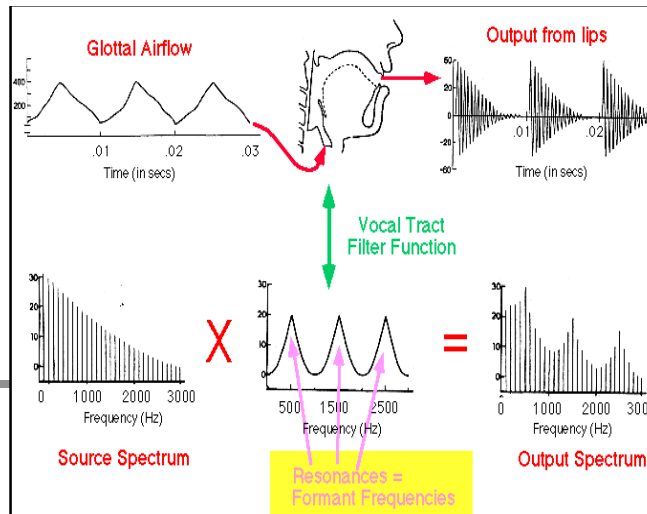
- 주파수(frequency: F)
  - 1초 동안 얼마나 자주 반복되는가를 나타내는 횟수
    - F=1/주기(p), F=1/0.01=100Hz
  - 복합파 (complex wave) ; 푸리에(Fourier, 19C 프랑스 수학자)
    - 서로 다른 주파수와 진폭, 위상을 가진 정현파들의 합으로 구성됨 p. 72(그림 3-3)
      - 주기파, 비주기파
- 스펙트럼(spectrum)
  - 음향학적인 신호를 주파수와 진폭으로 나타내는 것
- 위상(phase)
  - 두 개의 신호 사이의 타이밍 관계(timing relation)를 나타냄
    - 예) 타코마 다리, 공명 실험(스펀지)

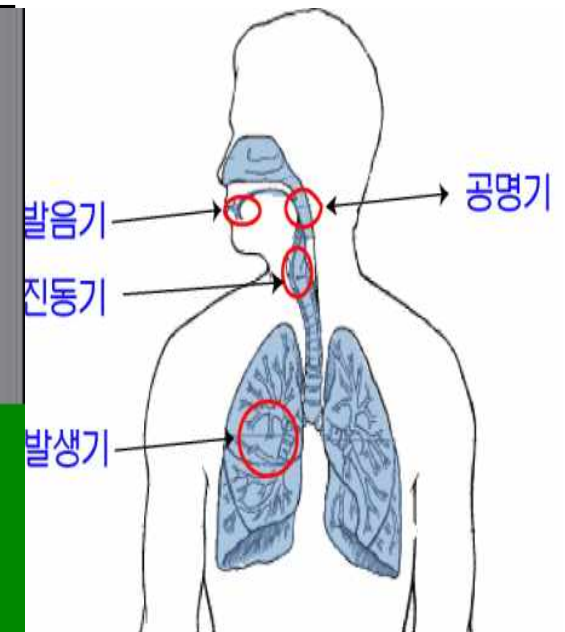
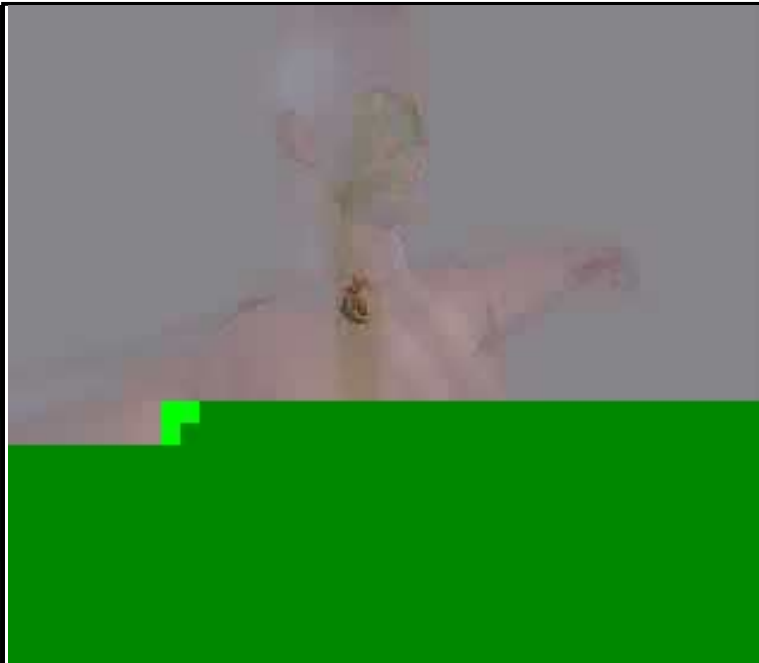
ㄱ)

sin θ	
0°	0
45°	0.71
90°	1
135°	0.71
180°	0
225°	−0.71
270°	−1
315°	−0.71
360°	0

ㄴ) ㄷ)

ㄱ)은 사인값을, ㄴ)은 90°의 위상차를 갖는 경우를, ㄷ)은 180°의 위상차를 갖는 경우를 각각 보인 것이다.

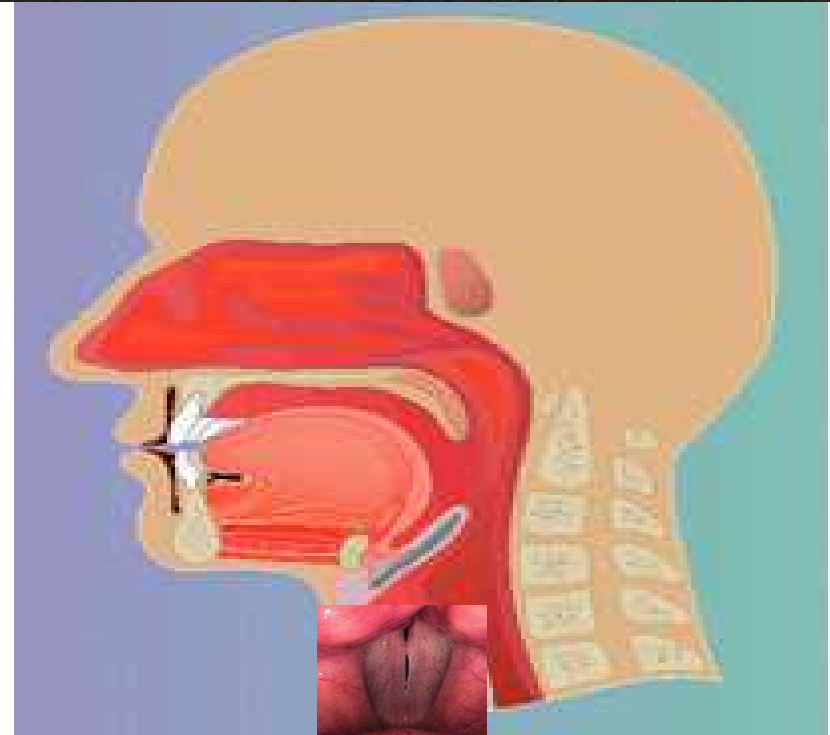
ㄷ)

ㄹ)

3) 공명과 포먼트

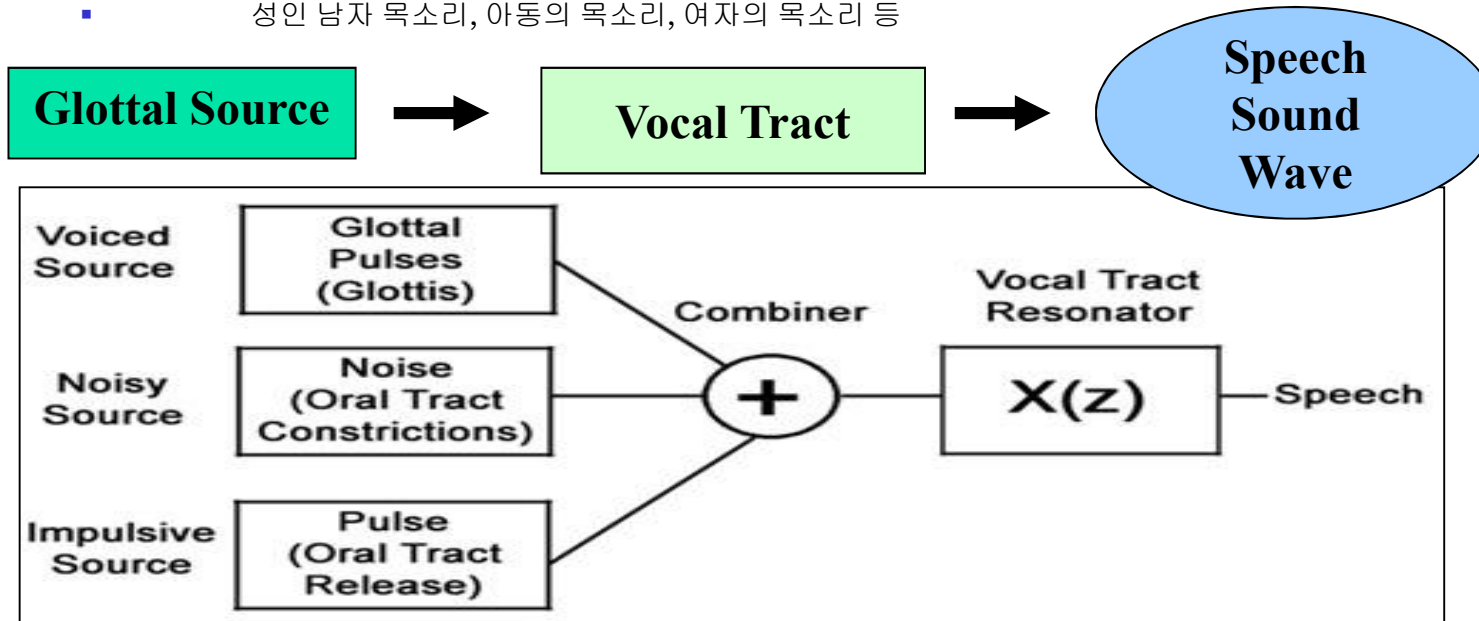
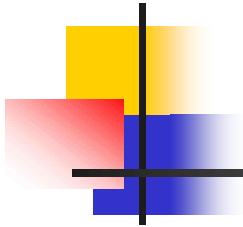
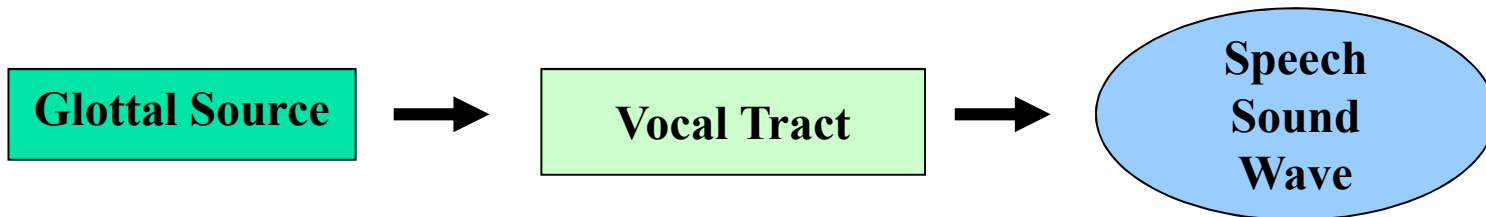
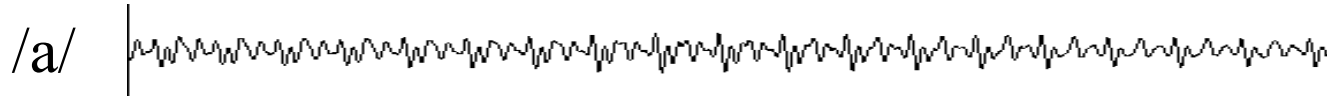
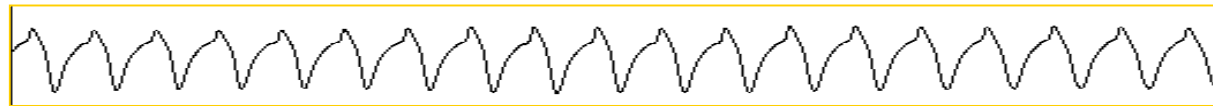
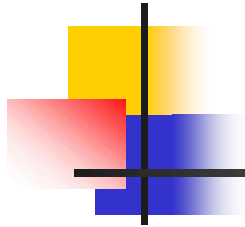
- 공명(**resonace**)
  - 성대의 떨림에 의해 생성되는 소리(울리는 소리)
    - 포먼트 주파수(formant frequency): 스펙트럼에서 뾰족한 봉우리에 해당하는 주파수들

- 포먼트(**formant**) ; 공명주파수 ; 음형대
  - 스펙트럼에서 뾰족한 봉우리에 해당하는 주파수
    - F1(제1포먼트), F2, F3, F4
  - 음성의 운율적 정보를 결정하는 음도와 더불어 음성의 특성을 나타냄
    - 음성인식, 화자인식, 음성분석 및 합성을 위한 측정치가 됨

3. 말 산출과정

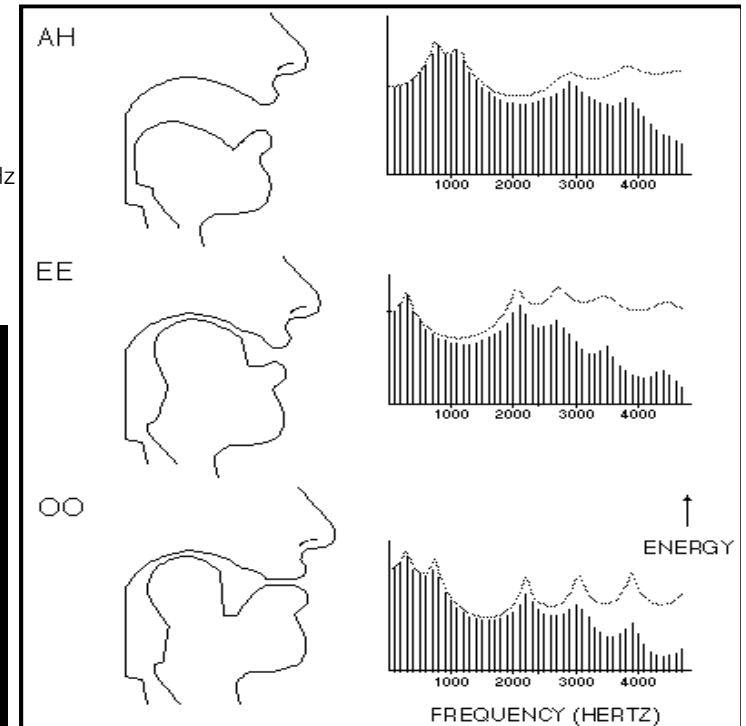
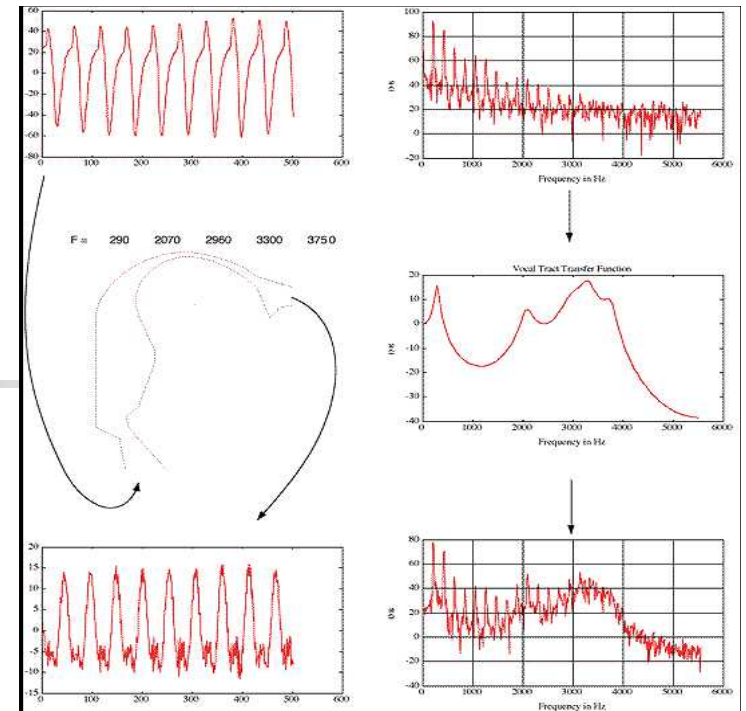
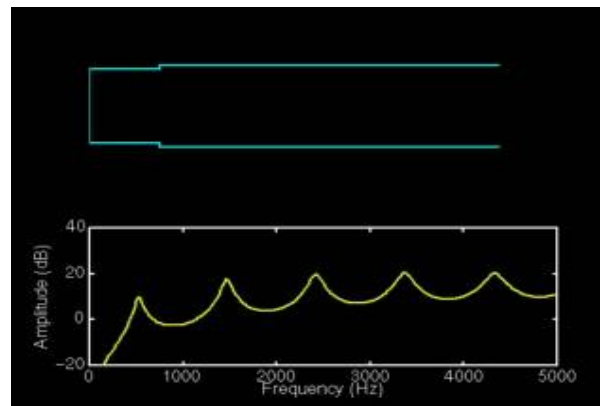
- 언어연쇄(**speech chain**)
  - 코 ➜ 비강 ➜ 비강인두(nasopharynx) ➜ 후두(성문) ➜ 기도 ➜ 폐 ➜ 폐포
    - 소리의 크기(60-65dB), 성문 밑의 공기압력(7-10cmH_2O)

1) 조음기관의 기능

- 여린 입천장, 혀, 입술, 턱 등
- 공명기관(F자 모양의 성도)
  - 성도
  - 비강
  - 구강
  - 인두강
  - 혀
  - 연구개 구조

2) 모음과 자음의 산출과정

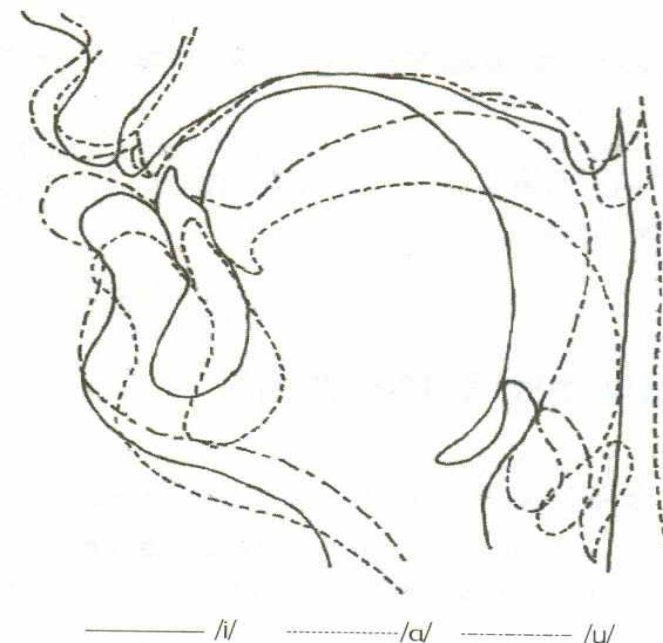
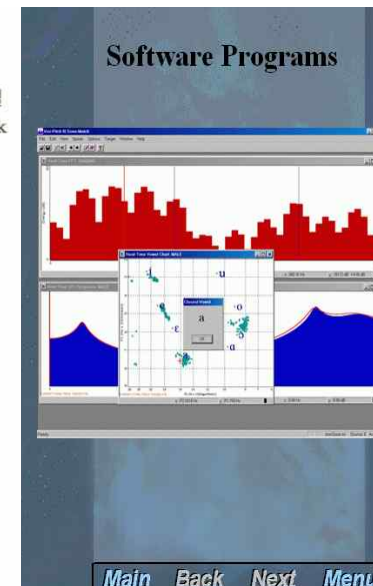
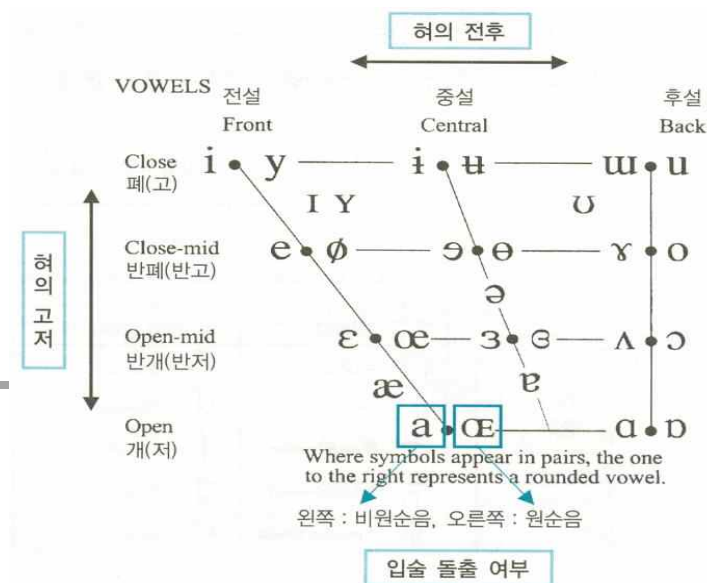
1. 음원필터이론(**Source-Filter theory;** *Fant, 1968*)
   - 소리에너지의 형성 과정 – 성대진동(vocal fold vibration)
   - 공명과정 – 말소리길(vocal tract; 성도)의 모양이 주요 역할을 함
     - 말소리를 만들어 내는 데 원료가 되는 것, 모음의 원료 : 성대의 진동음
     - 성도 : 울림통 역할
   - 성인 남자 목소리, 아동의 목소리, 여자의 목소리 등

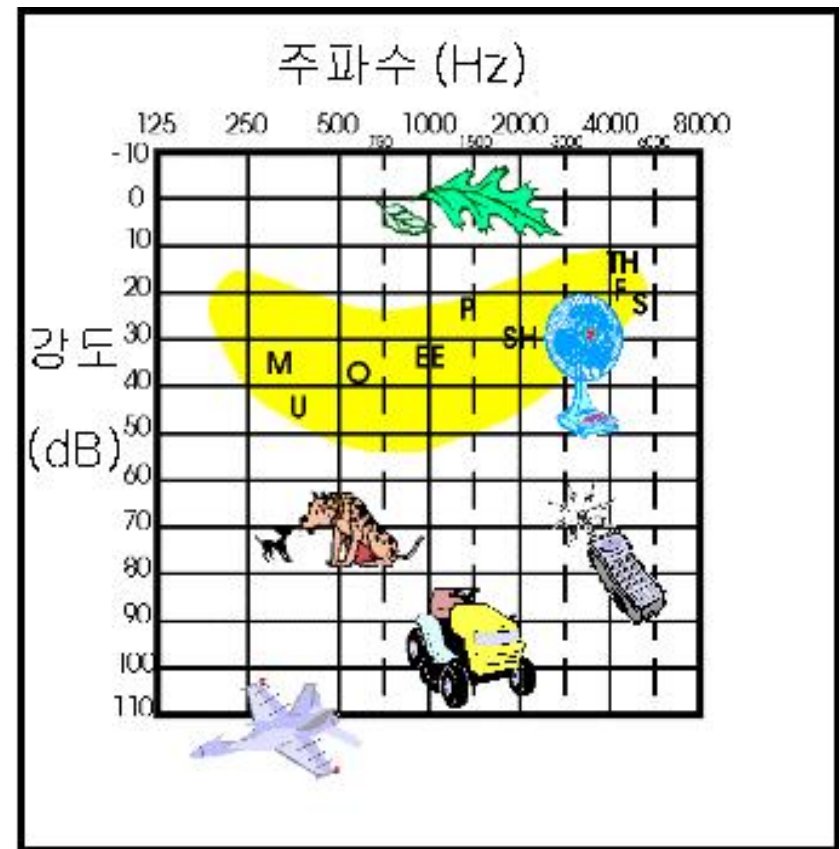
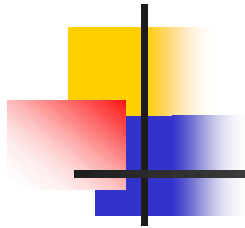
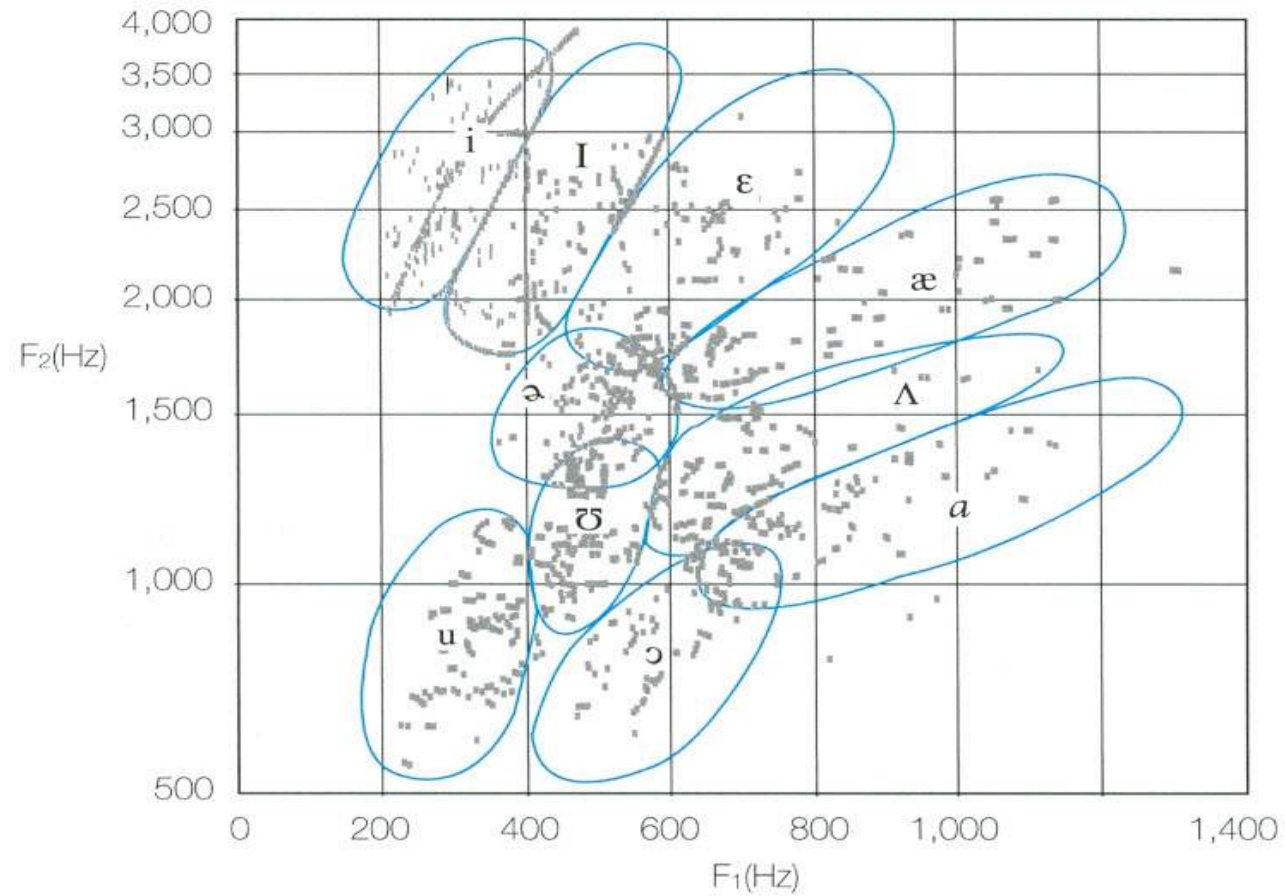
EBSHD

Glottal Source
/a/
/i/
/u/
Glottal Source
Vocal Tract
Speech
Sound
Wave

2. 모음의 산출과정

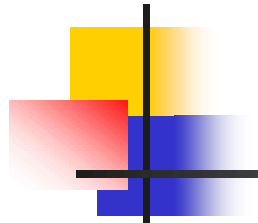
- $F_n=(2n-1)c/4l$
  - F : 공명주파수
  - Fn : n번째 공명주파수
  - C : 상수로서 음속 (약 340m/sec)
  - L : 성도의 길이(남; 17cm, 여; 15cm)
    - 예) 첫 번째 공명주파수
    - F1 : (2 ×1-1) ×34,000cm(4 ×17cm)=500Hz, F2 : 1,500Hz, F3 : 2,5000Hz
    - 관의 길이와 공명주파수는 반비례

Demo

3. 모음의 음향학적 특성

- 조음기관 위치와 포먼트
  - 성도의 길이가 길수록 모든 포먼트의 수치는 낮아진다.
  - 인후부가 좁아질수록 제1포먼트는 증가하고, 제2포먼트는 낮아진다(/아/).
  - 구강의 전방부가 좁아질수록 제1포먼트는 낮아지고, 제2포면트는 높아진다(/이/).
  - 턱이 많이 벌어질수록 제1포먼트는 증가한다.

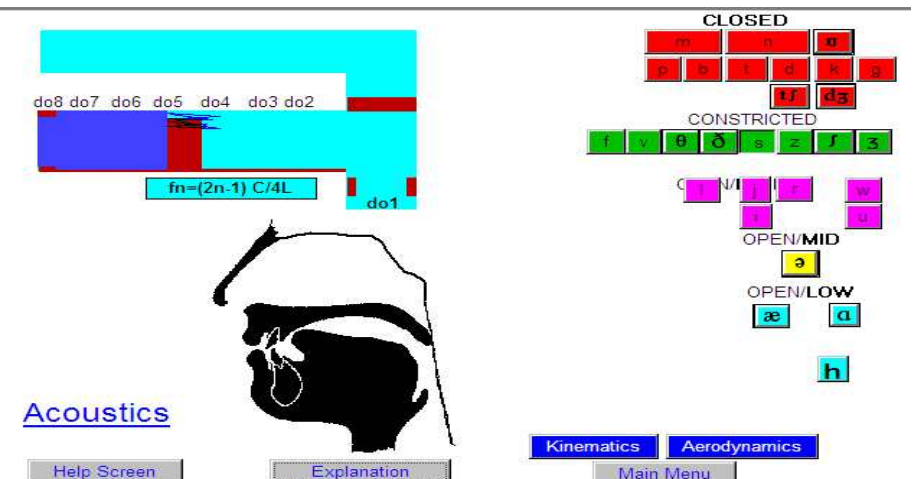
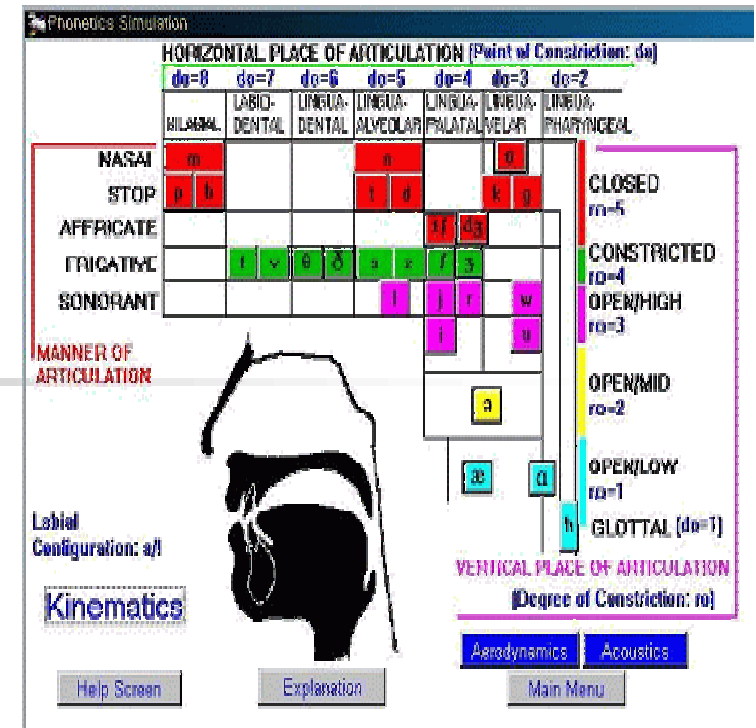
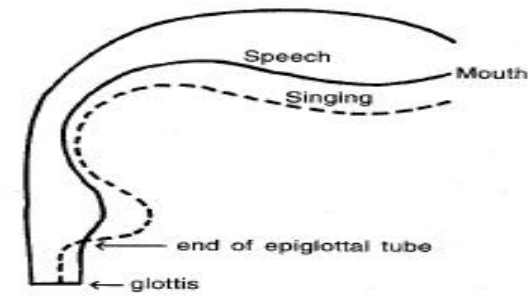
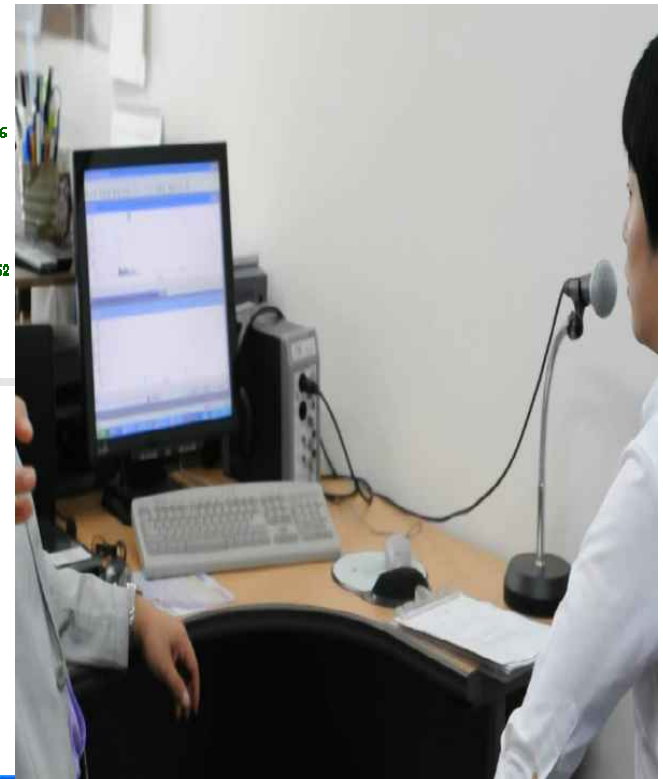
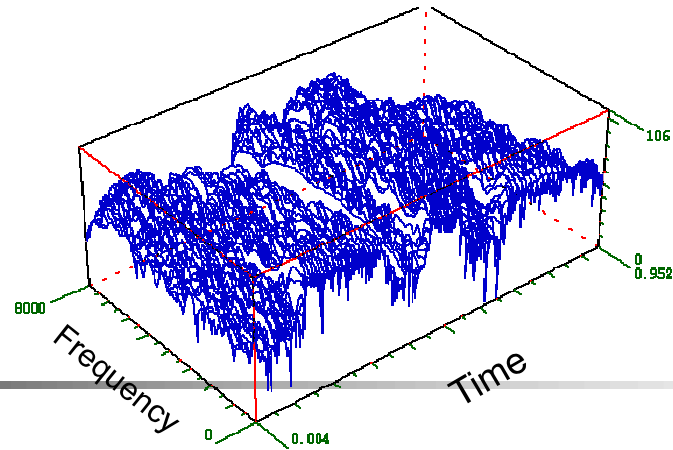
- **/a/ : F1은 높고 , F2는 낮다.**
- **/i/ : F1은 낮고 , F2는 높다.**
- **/u/ : F1과 F2 모두 낮다(대략 1,000Hz 아래에 분포함)**

Ling's 6 sound test

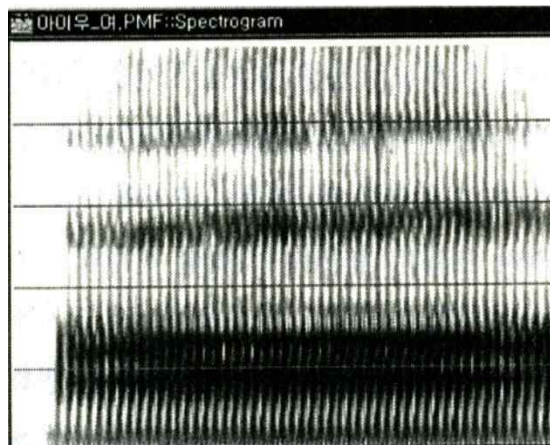
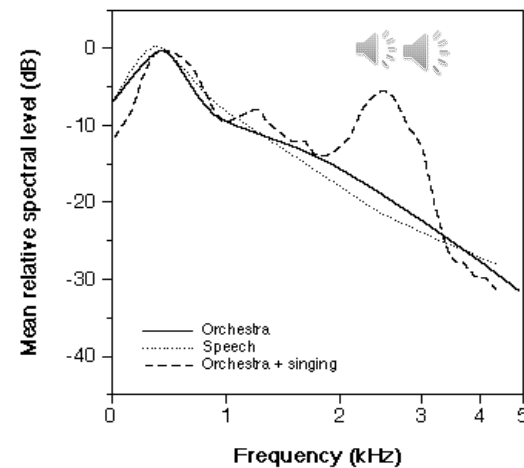
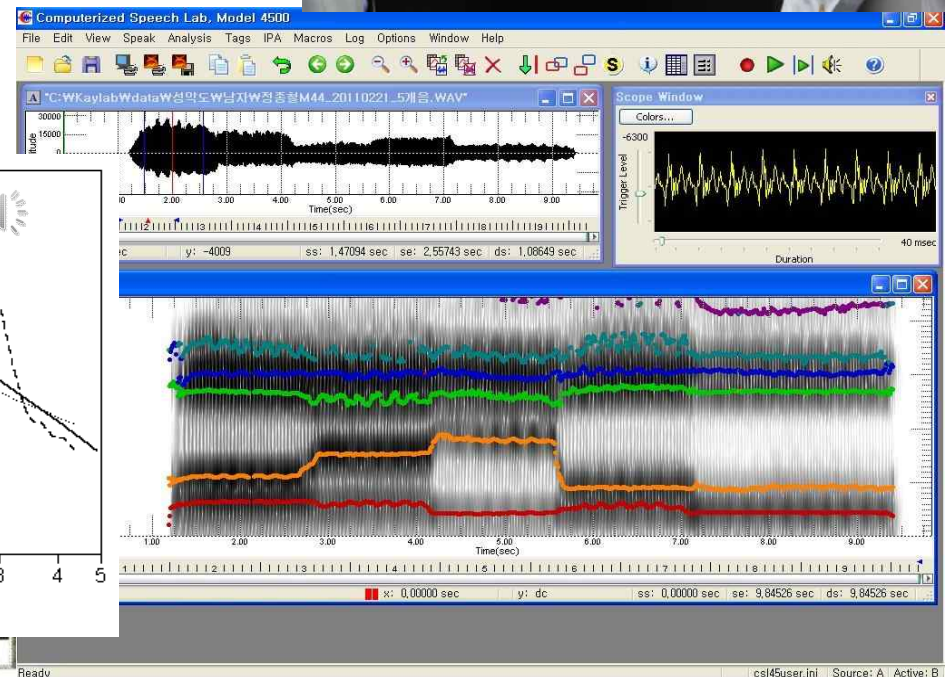
	성도공명주파수(Formants)	
	제1공명주파수	제2공명주파수
/m/	250 Hz	
/u/	300 Hz	900 Hz
/a/	700 Hz	1,200 Hz
/i/	300 Hz	2,500 Hz
/sh/	2,000 Hz – 4,000 Hz	
/s/	4,000 Hz – 6,000 Hz	

- 개별 모음의 모음공간

[그림 9-31] 개별 모음의 모음공간(vowel space)

출처: Peterson & Barney(1952: 182)에서 수정 인용

4. 자음의 산출과정

- 파열음, 파찰음, 마찰음, 비음 등

- **마찰음 (fricative)**
  - 기류가 통과하는 통로를 일정 정도 이상 좁혀서 난기류가 생성

- **폐쇄음 (정지음 stop; 파열음 plosive)**
  - 기류를 일정 정도의 시간 동안 아예 꽉 막았다가 한꺼번에 순간적으로 열어서 소리를 내는 것
  - 일정 시간 동안 성도의 일부를 조음기관을 이용하여 막았다가 터트리면서 내는 소리

5. 자음의 음향학적 특징

- **스펙트로그램(spectrogram; SPG)**
  - 음향 신호를 주파수, 강도, 시간으로 분석하여 얻어진 그림
  - 주파수와 진폭의 시간에 따른 변화를 보여주는 삼차원적인 그림
    - **X축 :** 시간
    - **Y축 :** 주파수
    - **Z축 :** 진폭

[단독] 보성 연쇄살인사건, 1.2초의 목소리로 범인 잡았다

2007. 08

[쿠키 사회] 119 신고전화에 불과 1.2초 동안 기록된 음성 흔적이 대학 연구팀의 치밀한 분석에 의해 범인의 목소리로 입증됐다.

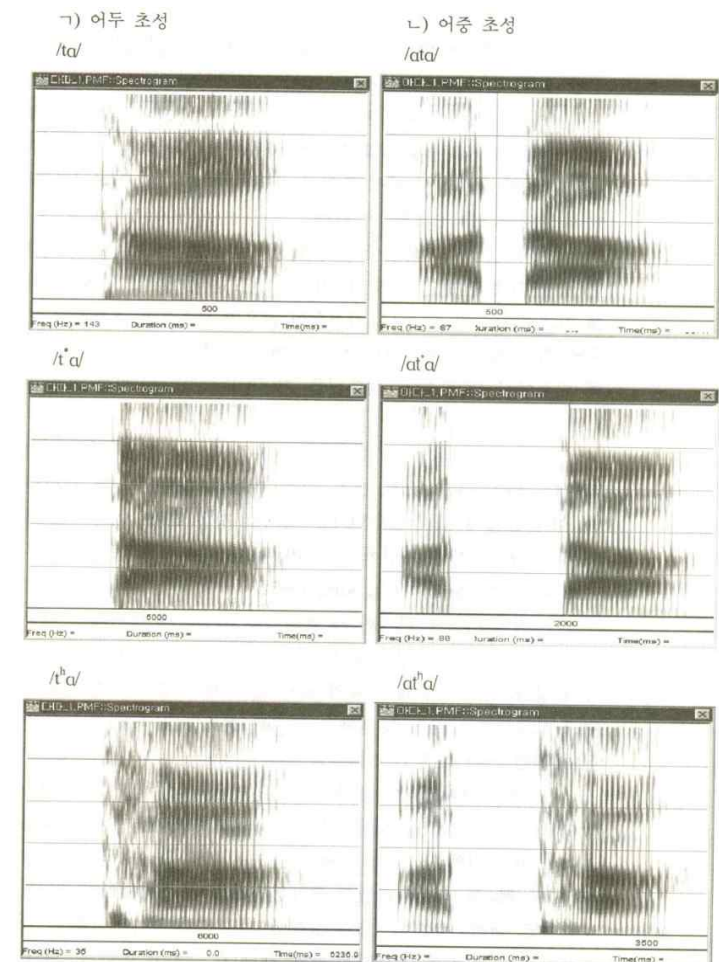
숭실대 소리공학연구소 배명진 교수 연구팀은 지난 8월 전남 보성에서 남녀 관광객 4명을 연쇄살해한 혐의를 받고 있는 어부 오모(70)씨의 목소리와, 피해자의 119 통화기록에 남은 범인의 목소리를 비교·분석해 오씨의 범행을 입증하는 결정적인 증거를 찾아냈다고 23일 밝혔다. 보성 연쇄살인 사건은 엽기적 범행으로 사회에 상당한 충격을 줬지만 검찰은 지금까지 오씨의 진술 이외에는 목격자나 마땅한 증거자료를 찾지 못해 수사에 어려움을 겪어왔다.

이 사건으로 숨진 여대생 C(20)씨는 납치 과정에서 4차례 119에 전화를 했으며, 이중 4차 통화에 "어디서 무전하니"라는 범인 음성이 고작 1.2초 동안 녹음돼있었다. 목소리 분석으로 본인 여부를 가리기 위해서는 최소 10분의 분량이 필요해 국립과학수사연구소 조차 119 통화 기록 속의 목소리가 오씨라는 사실을 밝혀내지 못했다.

그러나 배 교수팀은 나이가 들면 청력과 발성이 점차 약해져서 소리의 음 대역폭이 줄어드는 성문(聲紋)특성(목소리의 분위기 등 개인 특유의 정보)이 나타난다는 사실에 착안해 목소리 스펙트럼을 비교했다. 그 결과 119 통화에 남겨진 음성은 분석 스펙트럼의 300Hz 근방에서 소리성분(크기, 높낮이, 리듬 등)이 집중적으로 나타났으며, 스펙트럼의 주파수가 점차 감소하는 성문특성을 보였다. 그리고 이 특성은 검찰이 사건 발생 뒤 오씨에게서 녹음한 "어따 전화하고 있어"라는 목소리의 분석 내용과 정확하게 일치했다.

전화기록에 남은 음성이 오씨의 목소리라는 사실이 확인된 것이다. 이와함께 배 교수팀은 119 통화에 녹음된 선박의 엔진소리가 경찰이 나중에 녹음한 오씨 선박의 소리와 일치한다는 점도 밝혀냈다. 배 교수는 "비록 피해자는 유명을 달리했지만 그가 남긴 1.2초의 목소리가 범인을 규명할 강력한 증거로 작용했다"고 말했다. 국민일보 쿠키뉴스 노용택 기자

- **VOT**(Voice Onset Time; 발성시작시간)
  - 묵음구간, 기식화구간, 전이(transition) 구간, 안정구간 등
  - 구강협착(oral constriction)
  - 감폭현상(damping)

- 폐쇄음이 개방한 후에 후행하는 모음을 위해 성대가 진동하는 시간 사이의 간격
  - 기식성의 정도와 비례
  - 기식성이 큰 폐쇄음일수록 폐쇄기간 동안 성대가 멀리 떨어져 있는 상태로 조음됨
  - 경음<평음<기식음

[그림 8-15] 발성 유형에 따른 우리말 치경 폐쇄음의 음향적 차이

4. 말 지각과정

1. 모국어자기장 이론(native magnetic theory)
   - 젖먹이가 자신의 모국어에 지속적으로써 모국어 말 지각에 필
     속 발달하는 반면, 모국어 말 지각에 중요한 역할을 하지 못하는

2. 말 지각 운동이론(motor theory of speech p
   - 청자는 조음동작을 내적으로 모델링함으로써 상대방의 말

3. 정보처리과정(information processing)
   - 하상정보처리이론(bottom-up)
     - 말소리에 과한 음향학적 정보가 말 지각에 커다란 영향을 미침
   - 상하정보처리이론(top-down)
     - 인지적, 언어학적인 수준에서 이루어지는 가설검증과정

Cf) 칵테일파티 효과(cocktail party effect) ; bottom-up

- 인간은 원하는 소리에 집중하여 그 소리만 듣고 다른 소리를 구별할 수 있는 능력

- 사람은 말소리에 대한 원래의 음향학적 정보가 음운환경, 말하는 사람, 말 속도
- 변인에 의해서 변화되어도 전과 동일하게 말 지각을 하려 하기 때문이라는 것
- [Top-down이론에서 설명]